Opening access and closing risk: delivering the mandate for e-theses deposit

Wendy White, Institutional Repository Manager and Faculty Librarian Law, Arts and Social Science

University of Southampton Libraries, Hartley Library, University of Southampton, Highfield, Southampton, SO17 1BJ, UK whw@soton.ac.uk

Opening access and closing risk: delivering the mandate for e-theses deposit

The University of Southampton has a well established Institutional Research Repository (e-Prints Soton) with an embedded culture of deposit for research material for many discipline areas.[1] This paper explores the recent mandate to intergrate the electronic deposit of PhD and MPhil theses into the existing submission processes from the perspective of the students, the Academic Schools and the Institutional Repository. It examines the complex workflow issues for 23 Schools and cross-disciplinary research centres, covering the full range of academic subjects and the embedding of the theses deposits into the work of the repository team. Technical benefits and challenges are explored, particularly where there is a link between research and teaching. The PhD is conceptually ideally placed to facilitate this and supporting links to library catalogues and virtual learning environments like Blackboard is a key development. As many students progress into the research environment the open access e-thesis deposit adds value by exposing their research early and the Repository supports their ongoing research with export functions to CVs, grant proposals and personal home pages. Projects such as CLADDIER[2] are exploring the potential of linking data to research, including theses, with the overall aim of total access to all the components of the research process. At the University of Southampton this holistic view of research has incorporated economic issues as well as socio-cultural developments. Failure to protect Intellectual Property is a key concern and planning the changes in institutional processes to support e-theses has necessitated a fully joined-up approach between administrators and researchers, legal and research support staff, the Library and the Learning and Teaching Enhancement Unit. Risk management has drawn on diverse expertise to develop a workable non-bureaucratic process with relevant safeguards. It is only with such a safety net that people will truly feel confident about engaging fully with open access.

Open Access

The University of Southampton has a long standing commitment to opening out access to its research outputs. Initially the impetus for change came from within the research community in the Department for Electronics and Computer Science, now the School of Electronics and Computer Science (ECS). The creation of the ECS e-Prints service in 2000 encompassed all aspects of research, including theses, and the service currently gives access to 72 PhD theses, some of which have a long standing place at the top end of the viewing and download charts. The University wide research repository, e-Prints Soton, started at the end of 2003 and quickly moved from a JISC funded pilot to a fully embedded service.[3] This has a pre-mandate total of 180 theses that have been deposited as part of the regular depositing culture, excluding any retrospectively digitized theses.

The philosophy supporting the development of the repositories is to capture a public representation of the complete research output of the University and widen access to the full

[1] University of Southampton Research Repository, e-Prints Soton http://eprints.soton.ac.uk/ [accessed on 14/2/07] with a total of 22885 records, approximately 15% with full text/image. Metadata only records have been added to support the UK Research Assessment exercise but the proportion of records deposited with full text is increasing.

[2] The JISC funded CLADDIER project in which the University of Southampton is a partner, http://claddier.badc.ac.uk/trac [accessed on 14/2/07]

[3] Simpson, Pauline and Hey, Jessie M.N. (2005) TARDis: From project to embedded Institutional Repository, ALISS Quarterly, 1, (1), 12-14. http://eprints.soton.ac.uk/19122/ [accessed on 14/5/07]

content of the research where it is legal to do so. Theses are a core constituent of the research landscape and have the added cultural legacy of public engagement. In many European countries candidates undertake a public defense of their work. In the UK it is a requirement to make your thesis accessible in the British Library and usually your home institution unless there are proven matters of confidentiality. There is the possibility of embargo to protect IPR but it is expected that after three years the thesis will be made available to the public. The benefits and issues surrounding electronic access to theses relate to the medium, but the open access message should remain the same.

Engagement with e-thesis deposit not only helps display the full range of research but moves us a step further from research outputs as static individual deposits to the notion of the research journey. There has been much discussion about the life and half-life of individual papers as statistical analysis through metrics improves and it is easy to lose sight of the whole picture. The thesis is important because it marks the start of a journey, which may have few interesting diversions along the way and take in some notable landmarks. This is all the more reason to make sure that you create a map of progress as you go along. At Southampton we are interested in investigating further the range of activity that relates to this journey. New projects are underway to look at the deposit of cultural objects in the arts (KULTUR)[4] and small scale orphan datasets in the social sciences (DataShare)[5]. The exisiting CLADDIER project primarily deals with larger scale datasets, but a recent briefing paper outlines key issues which will also be relevent to data with theses.[6] The next step will be use these projects to help to move from the deposit and digitisation of predominantly print thesis to embrace the potential of the electronic environment. It may be possible to enhance both the research options and the richness of access. “Deep access” to the research could include data, visual and aural representations, links to supporting material such as lab notebooks or collaborative activity. This is an exciting prospect where work with theses could be well placed to lead the way.

To achieve progress it will be important to garner the enthusiasm of early career researchers. Mandating deposit will facilitate an electronic replication of the hard copy world. To build on this it will be important to combine mandates with research-led engagement to bring forth innovative exemplars. The role and impact of early career researchers is much discussed in the UK at the moment as institutions move towards their final Research Assessment Exercise (RAE) submissions. Early visibility of their work is a key incentive and there is an argument that the richer the picture the more likely it is that their work will be noticed and cited.

It is important that the model for the implementation of open access to e-theses embraces the potential of the cultural and technical environment, as well as addresses the practical and legal requirements. As we seek to continually enhance open access there is a cycle of activity that manages risk and embeds practice, but through collaboration and innovation consistently feeds through improvements that enrich the research landscape.

---

[4] KULTUR summary, JISC website http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/repositories_sue/kultur.aspx [accessed on 14/5/07]

[5] DISC-UK Datashare summary, JISC website http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/repositories_sue/datashare.aspx [accessed on 14/5/07]

[6] CLADDIER briefing on data publication, http://claddier.badc.ac.uk/trac/attachment/wiki/wp10-mtg/CLADDIER-datapub-briefing-20070514.pdf [accessed on 14/5/2007]

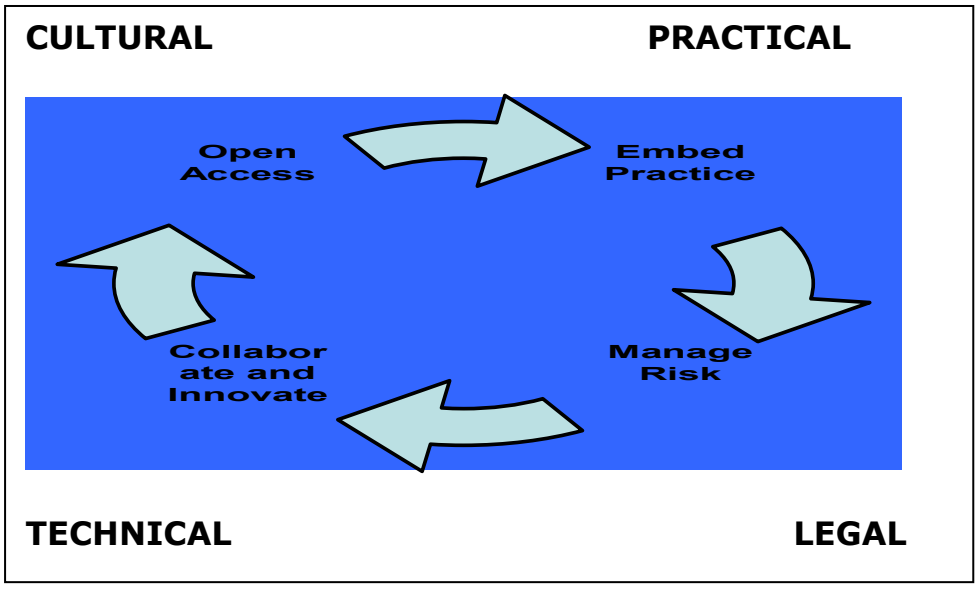

Figure 1: E-thesis Implementation Cycle

Embedding Practice

In Southampton we have a much used advocacy phrase "one record for many purposes" to promote the timesaving and added value benefits of deposit and eradicate any tendency to linger on the mundane aspects of filling in the metadata. One of the challenges for the deposit of theses is the fact that a lot of the embedded e-Prints Soton services, the elements that help embed the repository as a core University service, are of little incentive to someone who may well be about to leave and go to another institution. A streaming service for Schools or for individual staff can automatically populate School webpages or personal pages, an added incentive to put in the full outputs so they are downloadable not only through Google but through your personal showcase page. The research management tools are not much of a draw. Our current RAE reporting tools will be modified to include thesis reporting tools for administrators. Each School will be able to see how many thesis have been deposited by timescale and keep tabs on when embargoed theses are due to be automatically made public. There is also the issue that by the time the final version with all the corrections is completed the author may have started their new job and be halfway round the world. There are many benefits for the author to deposit their thesis but they are personal pluses and less likely to stem from institutional support structures and cultural practice.

There has been a lot of recent debate about the merits of mandates for deposit and whilst mandates are useful, it possible that they will only be as useful as the tools that you have to enforce them or incentives to encourage compliance. Theses have a ready made mandate built into the existing processes of most institutions. If the student does not deposit a final print copy in the Library they do not graduate. If you transfer this requirement to the electronic copy then the incentive is overwhelming and this is an institutional context which can compensate for the fact that the student may not benefit from the whole range of institutional benefits, although the generic open access benefits are compelling and there is an institutional role in promoting this broader cultural shift. New researchers can be encouraged to think of repositories as a globally linked set of services. Anyone who has a thesis and papers deposited in the research repository at Southampton can stream this information to any personal page or export information to a CV or grant proposal. They need not see this as an institutionally specific service. Institutional repositories can play a role in both promoting institutional identity and fostering collaborative activity, a network of institutions-without-walls.

From the perspective of repository development theses can be seen as deposit development leaders. As theses are submitted regularly throughout the year they encourage the institution as a whole to embrace a regular pattern of deposit. Repositories that lurch from one large upload to the next with stagnant weeks in-between are not yet on a firm foundation. A strong thesis culture can help develop daily activity whilst also encouraging new researchers to engage with repositories as a matter of course.
Theses are also well placed conceptually to facilitate links between research and teaching. Hard copy theses internally are not ideal for teaching purposes as in a large class the logisitcs of access become problematic and there is also the risk of damage. Access to e-theses can encourage the use of new research with undergraduates. Clear signposting and links to the full text through virtual learning environments and/or the library catalogue to supplement the exposure through the respository can make a significant impact. At Southampton we are encouraging links through our Blackboard VLE and embedding links to theses in the repository as part of our cataloguing process. We will be using the mandate to help review the technical processes and see if there are any cross-process innovations.

Managing Risk

There has been a very positive reception to mandating e-theses deposit at the University of Southampton with staff and students displaying a keenness to embrace the benefits of open electronic access. The idea was originally taken to the Graduate Schools Network for approval. Each Faculty has a Graduate School to promote training, support and opportunities for graduate students and they were happy to endorse the proposal. The University Research Policy Committee, chaired by the Deputy Vice Chancellor for Research, has now also approved the proposal. However in all cases the key issue for discussion has been the management of risk. The solutions need to be robust but not so administratively burdensome that it will encourage non-compliance. If the risk management model is itself to complex at best there may be reduced rates of deposit, at worst there may be people actively disengaging with or flouting key parts of the process that then increases not decreases risk.
Then there is the question of risk to whom? To the student if someone plagiarizes or misuses the work, to the institution if the terms of a sponsorship contract are breached or loss of income through lack of protection of Intellectual Property, to the supervisor of the thesis who feels responsible for the welfare of the student and the standing of the research group, to the Institutional Repository Manager who feels responsible for setting policies? The challenge is to take account of all these needs in a practical yet effective way and embed the process into the diverse cultural and administrative practices of 23 Schools of study. A framework was already in place for those Schools who had started to deposit theses and this was further developed by an E-Theses Working Group, drawn from academic representation from each Faculty. With a devolved University structure it would be easy for policy and practice to become fragmented across the Schools, therefore enthusiastic early adopters help strengthen the case for a University-wide mandate and process.
On the recommendation of the Working Group a number of improvements and new developments are being implemented. A “Permission to Deposit E-Thesis” form now includes not only the student’s signature but also the supervisor/s signature to ensure that no contracts are breached and to provide advice on the correct level of any required embargoes. For the student there is enhanced contextual information and help on format selection and PDF conversion. Part of the bigger picture is to link these changes to all the relevant sources of information from the formal University Regulations and Codes of Practice to Research Handbooks and School webpages. There is also a training need as researchers require more knowledge of file conversion and manipulation and as School administrators become more involved with institutional repository services. These skills should be seen as an essential part of the research toolbox and can be embedded in the existing research skills programmes required by the funding councils. A new stage will be added to the receipting process so that

when the administrators receive the thesis the student signs a simple statement to declare that it is the final version, including all corrections, at the same time they receive their receipt of deposit. The administrators will have all the documentation with each thesis and will manage the upload process.

Many of the developments to support the deposit of e-thesis do not stand alone but need to be integrated into the overall repository workflow and the risk management tools need to be part of the overall risk management strategy for the repository as a whole. A standard copyright statement sits as the front page of each thesis which will be automatically added on the fly to all theses that are uploaded. This is in addition to a generic click-through download license agreement for all items in the repository. User feedback has been important in helping refine these risk management tools. It is important to set a solid foundation through consistent practice and Southampton like many insititutions is focussing on getting this right, but there is a danger that consistency could be at the expense of flexibility. The author of a PhD may actively desire to set terms of use which are broader than the standardized approach and one of the key issues for future development is the added complexity of embracing an approach like Creative Commons as an intergrated approach, rather than a one off or an optional extra.[7]

All the metadata and files will be checked by the repository editorial team before they are made visible. The repository structure is set up to embrace the totality of the research output and the workflow for e-thesis deposit is just one of the areas which will test the scalability of the service. The workflow starts with the author and ends with the reader, but there are other key staff involved around the University, summarised in this workflow and communication diagram. The key action route from the researcher to the user is marked, along with all the other communication routes and agents.

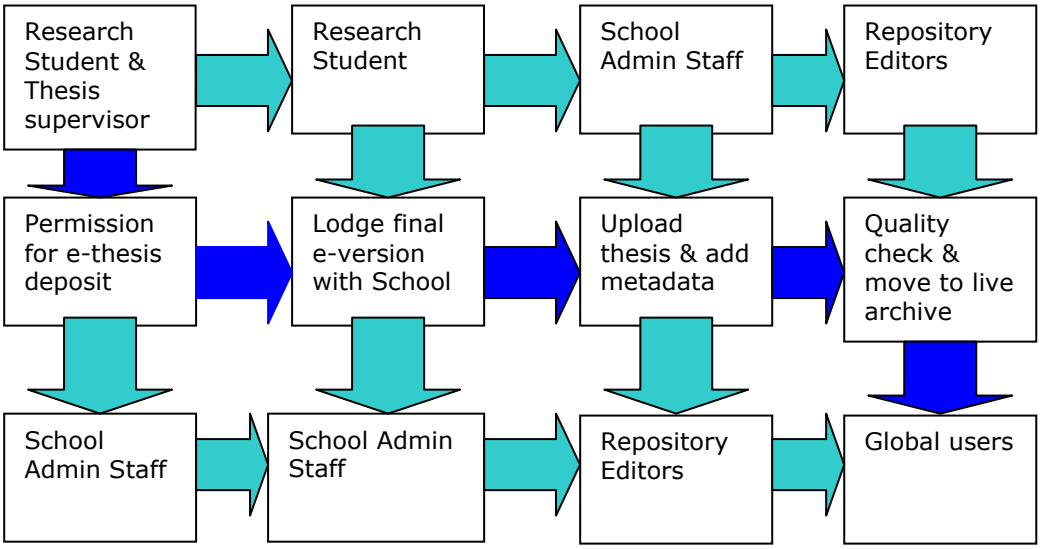

Figure 2: E-theses workflow and communication

Collaborative Innovation

[7] Klang, Matthias (2006) Disruptive technology: effects of technology regulation on democracy, Gothenberg Studies in Informatics, Report 36. http://www.ituniv.se/~klang/web/thesis/] [accessed on 14/5/07] is an example of personal inititative-led use of creative commons.
Fitzgerald, Brian Creative Commons (CC) licences, open access and electronic theses: opportunities and challenges http://adt.caul.edu.au/etd2005/papers/126Fitzgerald.pdf [accessed on 14/5/07]outlines some of the issues.

Little development would have been possible without collaboration, both internally and externally. Key internal participators have been the Deputy Vice-Chancellor for Research, Deputy Deans Research, the Faculty Graduate Schools, School Postgraduate Coordinators, School Managers, research administrators, the Learning and Teaching Enhancement Unit which deals with quality enhancement and assurance, and the E-Prints Steering group made up of representatives from the Library, Information Systems and Services and the School of Electronics and Computer Science. These collaborations and partnerships have been both formal and informal and crucial to the testing and development of principles and practice. The fact that they span the full range of roles within the University is an essential platform for progress. One of the key collaborations has been with our Legal Services team and with the Research Support Office who have responsibility for industrial contracts. Their advice has been extremely valuable, and as the legal aspects of research and teaching increase the need for Universities to retain and expand their legal team will be paramount. This may be a challenge in these times of competing call on resources.

The feedback loop with our core user groups, both in terms of those depositing material and those using deposited material, is a cycle of activity that could potentially feed valuable information into collaborative developments. Where we are pushing forward together in relatively new areas sharing what we learn from user feedback will be an important constituent of the development cycle. Creating opportunities for staff to learn from the experience of colleagues at other institutions in a timely, effective and resource-light way is a challenge.

The University of Southampton is proposing to run an electronic thesis deposit in parallel with a print deposit for 2 years but with a view to moving to electronic only deposit. This will give us time to build on on-going work to develop preservation policies and tools through projects such as PRESERV[8], whilst at the same time develop an institutional preservation strategy. Many Universities are in the same position amassing a bank of electronic assets, including born-digital material, in advance of an overall strategy. This is an area where success can only be collaborative, linking expertise across a range of institutions.

[8] PRESERV: Preservation E-Prints Services, project website, http://preserv.eprints.org/ [accessed on 14/5/07]